

TEXTUAL ORIGINS OF THE INDEFINITE ARTICLE IN SWEDISH

DOMINIKA SKRZYPEK

Adam Mickiewicz University in Poznań



ABSTRACT. The present paper studies the earliest stages of the grammaticalization of indefinite article in Old Swedish. The study is based on a corpus of Old Swedish texts and uses the model of grammaticalization as proposed by Heine 1997. The article *en*, etymologically related to the numeral 'one', is first used to mark new and salient discourse-referents and its primary function is cataphoric. However, *en* only fulfills this function when occurring in a sentence-initial subject NP. In the course of the grammaticalization, neither the sentence-initial position nor the subject function of the NP are required to present new and salient discourse referents.

1. INTRODUCTION

The Modern Swedish (MS) indefinite article EN takes its origin in the numeral 'one' and is formally identical with it (unlike e.g. English where the indefinite *a* has a different form from *one*, with which it shares a common source, *ān*), apart from the phonological realization since the numeral is normally stressed but the article never is. In its indefinite article use, *en* is lacking in the oldest extant sources, runic inscriptions, as well as in the oldest manuscripts containing Old Swedish laws (from the 13th and 14th century). Its formation can be studied in somewhat younger prosaic texts and follows the grammaticalization stages of the indefinite article as outlined in Givón 1981 and Heine 1997 (see 2). It is inflected for gender, case and number and will here be referred to as EN in all its uses, numeral and article ones alike.

At the time of the inception of this grammaticalization, Old Swedish has a grammaticalizing definite article, etymologically related to the demonstrative *hinn* 'yon', which is at this point used to mark mainly anaphoric text-referents,

but which has not yet reached all contexts available to a definite article in MS (see Skrzypek 2012, 78-155 for a detailed overview of the grammaticalization of the definite article in Swedish). In the texts it appears as a suffix¹ and will be here referred to as -IN.

The onset of grammaticalization of the indefinite article is the use of the numeral in the so-called presentative function, to mark new text-referents. The presentative function implies that the new text-referent is salient and persistent. Initially, EN cannot be used with text-referents which are dropped in later discourse. The presentative function has been identified as the first stage of the formation of the indefinite article in e.g. Heine 1997. At this stage of development other items, such as 'some' or 'such' can be used, there is a high level of competition between the forms; the one competing against EN in Old Swedish is *någon* 'some/any' (see 5). The study of the Old Swedish material suggests that the use of EN is presentative only under certain conditions, i.e. in conjunction with the function of the NP and its position in the clause (see 6). No such restrictions apply to forms other than EN, competing against it in the presentative function.

The aim of the present article is to outline the earliest stages of the grammaticalization of the indefinite article in Old Swedish (OS, 1225-1526). The study is based on a textual corpus of 180000 words, containing texts from the OS period (see 3). The development is seen as at least partly dependent on the parallel formation of the definite article, an interdependence which seems to be characteristic of the developments in other Indo-European languages (see 8). The formation of the indefinite article in Swedish has been treated before in Brandtler and Delsing 2010 and Skrzypek 2012. The first publication focuses on stage IV of the grammaticalization (see 2) and in particular on the use of EN as a non-specific marker. The latter is a comprehensive study of the grammaticalization of both definite and indefinite articles in Old Swedish. Here, I would like to concentrate on the earliest stages of development.

2. MODEL OF GRAMMATICALIZATION OF THE INDEFINITE ARTICLE

There are a number of linguistic items that may serve to introduce new text-referents into the discourse, apart from the indefinite article. They include pronouns such as 'such' or 'some' and adjectives such as 'certain' and could possibly serve as the source of the indefinites. However, it seems to be a

¹ Some authors, such as Syrett 2002, use the term clitic. Such classification relies on prosody (clitics do not change the tonal patterns of their hosts and the definite article in Swedish indeed does not influence the host). In other respects, however, -IN exhibits properties of a suffix, e.g. it can never be separated from its host and all inflectional forms (such as the genitive -s) are attached after the suffix.

universal tendency of the world's languages that the only item that grammaticalizes into the indefinite article is the numeral 'one' (Heine and Kuteva 2002:161ff).

The grammaticalization process whereby the numeral turns into an article can be divided into five stages (following Givón 1981 i Heine 1997).

- Stage I: the numeral 'one'
- Stage II: the presentative marker
- Stage III: the specific marker
- Stage IV: the nonspecific marker
- Stage V: the generalized article

Diagram 1. *Model of the grammaticalization of the indefinite article, after Heine (1997:72-3)*

The model builds on the observation that the indefinite article may be used to express both specific and non-specific reference. In some contexts the interpretation of the article may be specific ('a certain'), non-specific ('any') or opaque (specific or non-specific depending on the wider context), see examples 1-3.

- | | | |
|-----|--|--|
| (1) | Scarlett married a millionaire. <i>specific</i> | He was called Rhett. |
| (2) | Scarlett didn't marry a millionaire. <i>non-specific</i> | * He was called Rhett. |
| (3) | Scarlett wanted to marry a millionaire. <i>opaque</i> | He rejected her./She didn't find any . |

The examples above illustrate further that depending on the interpretation of the text-referent as specific or non-specific it may either be picked up in later discourse (specific) or not, blocking the anaphoric reference (non-specific).

As it turns out, the numeral 'one' does not occur in all these functions simultaneously. Its first article-like function is what Heine terms 'the presentative marker'. In this function 'one' is similar to the Modern English demonstrative *this*, as in the following example:

- (4) I was sitting on the bench when suddenly **this** man started screaming.

The function of the presentative marker is strictly textual: to introduce new and salient text-referents. The salience of the text-referents can be measured by their lifespan (Heim 1988), i.e. the number of further mentions of the text-referent in the discourse.

The lifespan is what keeps the presentative and specific markers apart: while the former initiates an anaphoric chain, the latter need not. In other words, the hearer will expect a text-referent introduced by means of a

presentative marker to be picked up in the later discourse while a text-referent introduced by means of a specific marker may be picked up, but need not be. This is particularly clear in the case of the English indefinite article, where two items, *an* and *sum*, are to begin with in competition in the presentative function. The variation between them seems to be quite systematic: *sum* serves to introduce new and salient text-referents, while *an* introduces less salient ones which are less likely to be picked up in later discourse (Hopper and Martin 1987). Hopper and Martin note further that there is a difference in terms of the position in which the two forms may appear: while *sum* is typically located at the beginning of the text, *an* can appear anywhere within it. We shall see in 6 that the location of *en* in Old Swedish changes in a similar way, as it grammaticalizes from a presentative to a specific marker.

The fourth stage of the grammaticalization of the indefinite article is a non-specific marker. A non-specific marker is used to indicate that the noun marked is not meant to refer. As has been observed in literature there are a number of contexts where unless the noun is definite it must be regarded as not only indefinite but also nonspecific (e.g. Barwise and Cooper 1981). These are so-called negative polarity contexts where the noun is found e.g. in the scope of negation, hypothetical conditional, in a comparison or after a verb of volition (e.g. *to want*). Finally, the generalized indefinite article can be used in generic contexts.

3. THE OLD SWEDISH SOURCES

The present study is based on a corpus comprising 17 OS texts (see Sources for details), representing three genres: legal prose, religious prose and narratives. Poetry was excluded from the study, as earlier research shows convincingly a significant influence of Middle Low German on the article use there (e.g. the preference for preposed definite instead of suffixed article, see Haskå 1972). The influence of foreign languages, mainly Latin and Low German, on other texts in the corpus at hand cannot be ruled out, since they are, with the exception of legal texts, translations. This may influence both article use and syntactic structures (positions of NPs in the sentence structure). A detailed comparison with the original texts would undoubtedly broaden our understanding of the development, it falls however beyond the frames of the present study.

The corpus is ca 180000 words long. The texts were excerpted in their electronic form (available at *Fornsvenska textbanken*, see Sources) by means of a concordance creator KWIC; the concordances were manually pruned and checked against edited versions of manuscripts (see Sources for a complete overview of the source texts).

It has become a common practice to subdivide the OS period into three subperiods: 1225-1375, 1375-1450 and 1450-1526 (see e.g. Hirvonen 1997, Håkansson 2008) rather than two, as proposed in textbooks (1225-1375 and 1375-1526). Since such subdivision allows for more precise study of changes in the development of many categories, definiteness included, this practice is followed here. From period 1 (1225-1375) the following texts were excerpted: AVL, YVL, OgL, DL, Bur and Pent; from period 2 (1375-1450) Bo, HML, Jart, SVM, KM, JB; from period 3 (1450-1526) PK, Did, Linc JB, Linc NR and Troj, which being written down in 1529 formally does not belong to the Old Swedish period, however, is linguistically treated as part of the OS text heritage.

4. THE METHOD

The main indication of grammaticalization of a form is a rise in its frequency. As the form reaches new stages of grammaticalization and spreads to new contexts, its frequency rises. Therefore, the simplest tool for measuring the level of grammaticalization is measuring the average number of instances of the form per 1000 words and the mean value for all the texts from one period. However, since there are great discrepancies between texts from the same period, partly depending on the genre of the text, it is also interesting to give the median results. The median is the numerical value separating the higher half of the data sample from the lower one; if the sample consists of an even number of data the median is the mean of the two middle values, if the sample consists of an uneven number of data, the median equals the middle value. This allows us to reduce the influence of the so-called outliers (the most extreme observations), whose values may depend on other factors (in the case of historical linguistics such as the dialect of the scribe, mistakes created by copying the text from an older version etc.). If there is a difference between the mean and the median, it suggests that there are significant discrepancies between the texts in terms of the frequency of the grammaticalizing element. If, however, the difference is small, it suggests that the genre differences and other factors do not influence the frequency of the form in different texts. Some differences are not to be avoided, since genre does play a role in the frequency of the articles. As studies of Modern Swedish show, legal prose tends to use more bare nouns than other genres and the results for articles are lower (Gunnarsson 1982). However, the genre-dependent variation is limited in MS.

In the present study, the general results, the mean and the median, are further fine-grained, when the function of the NP (subject, object, PP, other) and its position (sentence-initial, sentence-medial, sentence-final) are taken into account. This allows for a more precise description of the process,

showing that the initial rise in frequency may be limited to certain functions and positions.

Since the presentative function implies that the new text-referent be salient, we expect it to be picked up in later discourse. Therefore, a further, manual search was conducted on the material excerpted, concerned with referent tracking. Further mentions of the new text-referent were listed and counted. The results show how a number of further mentions of a new text-referent is dependent on the function of the EN-phrase as well as on its position in the sentence.

The method applied is of necessity statistical. The subtle variation in the use of the form in a given context can only be tested with native speaker information, which is unavailable. However, since grammaticalization is so closely connected with frequency, it seems to be a reliable tool for measuring its advancement.

5. ETYMOLOGIES

EN is a descendant of the Proto-Germanic **aina-* from ie. **oi-no* (Lat. *unus*, Greek dial. *oinós*). In the legal texts, which are the oldest texts in the material studied, EN occurs only in its original role of the numeral, which is usually clear from the context giving a contrasting numeral, as in (5).

| | | | | | | | |
|-----|-------------------|-------------|-------------------|--------------|------------------|----------------------|-----------------|
| (5) | Drukknæ drown | tuer two | baþir both | a on | enu EN | skipi ship | hvarghi each |
| | þærra they-GEN | ær is | annars other's | arve heir | | | (AVL ÆB:13) |

'If two men drown on one ship, they are each the other's heir'.

A form competing with EN in presenting new referents (but only non-specific ones) is *nāgon*, a relatively late formation which unlike EN does not trace back to a single Indo-European element. It was formed in the Ancient Nordic/Old Nordic time (Hellquist 1957:713) as a contraction of the phrase *ne wait ek hwariR*, literally 'not know I which'. In OS sources it appears as *nokorr* and will be here referred to as NGN. Since the competition concerns the non-specific contexts, it appears first when EN has reached stage IV of grammaticalization.

6. THE RESULTS

It is typical for a grammaticalizing item that its frequency rises in consecutive periods. So is the case with EN in OS as well.

| | Period I | Period II | Period III | MS ² |
|--------------------------------|----------|-----------|------------|-----------------|
| EN ³ per 1000 words | 3,82 | 8,7 | 9,56 | 23 |
| % | 0,382% | 0,87% | 0,956% | 2,3% |

Table 1. *Adnominal EN per 1000 words (Skrzypek 2012:156).*

The values given in Table 1 are the mean values for each period. They show how the frequency of EN rises gradually between period I and III. What they fail to show is that there are significant discrepancies between the texts in a given period, particularly in period II (only 2,51 in Bo, but as much as 22,6 in Jart). Since both texts represent same genre, religious prose, the difference is not genre-dependent. Instead, an explanation may lie in their overall structure: while Bo is a coherent narrative with a few main characters whose life we follow, Jart is a collection of short stories, each introducing new text-referents (by means of EN).

It is to be noted that the differences in frequencies of EN between texts from the same period are reduced as the grammaticalization progresses and the item becomes obligatory in more and more contexts. Table 2 below shows the median results for all periods as well as the mean results.

| | Period I | Period II | Period III | MS |
|--------|----------|-----------|------------|----|
| median | 3,2 | 5,8 | 8,64 | 22 |
| mean | 3,82 | 8,7 | 9,56 | 23 |

Table 2. *Median and mean values in periods I-III in Old Swedish*

We can see that while median results are lower than the mean in all periods, the distance between the two differs from period till period. It is largest in period II, but the two values start to coincide again in period III. How are these results to be interpreted?

In period I the distance between the mean and the median is relatively small. There are discrepancies between the frequencies in individual texts, mainly depended on the genre: legal texts show lower figures than religious prose. Moreover, all occurrences of EN in the legal texts represent stage I, the numeral 'one'. It is only in the religious prose that we find first instances representative of stage II, the presentative marker. Since the frequency of the latter is low in religious prose of period I (6,14 per 1000 words in Bur and 3,47 per 1000 words in Pent), no significant distance between the mean and the

² MS: Modern Swedish. Comparative MS values are based on three short stories (see Sources).

³ EN is used here to denote any occurrence of forms of *en* (including the inflected forms, e.g. *en-s en-GEN*). Thus, terms such as 'EN-subject' used throughout the text denote any subject which takes form of a noun phrase modified by any form of *en*.

median is shown in period I. However, a more noticeable distance between mean and median is found in period II (see Table 2). Here the differences in frequencies from one text to the other do not seem to depend on the genre. It must be noted that period II is the time when the grammaticalization of the indefinite article gains ground. It means that EN is no longer found in stage II contexts (presentative marker) only, but in stage III ones as well, with occasional occurrences in stage IV types of context (see Skrzypek 2012:174-184). However, not all texts in period II show same level of grammaticalization of EN. The most extreme values are found in two religious texts, Jart (22,6 per 1000 words) and Bo (2,51 per 1000 words). Although more contexts are available to it, the form is clearly not yet obligatory and the author has a choice whether to use it or not. The frequency of EN stabilizes in period III, when the relatively close values for the median and the average show that the frequency of EN is independent of the genre.

Why is this an important observation? One aspect of grammaticalization is the obligatorification of the form in certain contexts. Thus the use of the form is no longer only stylistically motivated and possibly genre-dependent, but is obligatory for grammatical reasons, i.e. the introduction of a new discourse-referent must always be marked with EN. Therefore, we expect the median values to coincide with the mean values, since even the outliers (texts with extreme values) will show same patterns of use of the form and therefore similar frequencies. Examples (6) and (7) illustrate how EN is missing from a PP context in a period 1 text, but is found in same type of context in a period III text. The contexts are similar in that they introduce a new but inimportant text-referent.

Pent (period 1, religious prose)

| | | | | | | | | | |
|-----|---------------------|-----------------|---------------------|-----------------|-------------|-------------------------------|---------------------|---------------------|-------------------|
| (6) | Nw now | kombir comes | ioseph Joseph | fram forward | til to | sinna his | brødhra brothers | oc and | æmskøt at once |
| | føra take | the they | han him | aff of | hans his | kledhom clothes | | oc and | |
| | kastado threw | han him | liffuandis alive | | ī in | vatnløsan waterless | | brwn well | æpthir after |
| | rubens Ruben-GEN | | radhe advice | | | | | | (Pent 178) |

‘Now Joseph comes forward to his brothers and they at once take his clothes off his back and threw him into a waterless well, still alive, as Ruben advised’.

Did (period 3, narrative)

| | | | | | | | |
|-----|-----------|-------------|---------------|---------|-----------------|------------------------|------------------------|
| (7) | han he | lag lied | lenge long | ī in | en EN | storan large | skogh forest |
|-----|-----------|-------------|---------------|---------|-----------------|------------------------|------------------------|

oc giorde sik ther et hws
and make self there EN house (Did 3)

'He stayed in a forest for a long time, and made himself a house there'.

However, it is not only the frequencies that are changing between periods I and III. A closer look at the functions of EN-phrases reveals that in different periods EN is used in different functions. These are here quoted as EN-subject (= an NP with EN in a function of a subject), EN-object (= an NP with EN in a function of a object) and EN-PP (= a PP governing an EN-phrase). An overview is given in Table 3. The results relate to the percentage of the type of EN-phrase in all EN-phrases in a given period.

| | Period I | Period II | Period III |
|-----------------------------------|----------|-----------|------------|
| EN-subject | | | |
| EN-object | 21,4% | 20,2% | 30,2% |
| EN-PP | 19,15% | 24,7% | 25,5% |
| Other (attribute, adverbial etc.) | 32,4% | 25,1% | 31,8% |
| sum | 100% | 100% | 100% |

Table 3. *Grammatical functions of EN-phrases in OS.*

The most significant difference is between the percentage of EN-subjects in periods I and II on the one hand and period III on the other. The values fall, while the values for EN-objects rise. The system evolves towards the MS one, where new discourse referents are typically introduced by means of EN-objects.

However, there is another striking difference between the subjects in periods I-III which the general results in Table 3 do not show. It has to do with the position of the EN-subject in the sentence: initial or non-initial. The results are given in Table 4 below.

| | Period I | Period II | Period III |
|-------------------------|----------|-----------|------------|
| Initial EN-subjects | 63,2% | 41,3% | 20,8% |
| Non-initial EN-subjects | 36,8% | 58,7% | 79,2% |
| sum | 100% | 100% | 100% |

Table 4. *Initial and non-initial EN-subjects in OS.*

As can be seen, the position of EN-subject changes significantly from period I to II. In period I, the EN-subjects can be mainly found at the beginning of the sentence, a position which in MS is only rarely occupied by indefinite NPs⁴. In MS the subject is often the theme of the sentence and as

⁴ It is not impossible to place an indefinite subject sentence-initially in MS, however, the NP must be anchored in earlier discourse, as in e.g. *Foten var värre. En lukt började kännas*

such known or identifiable to the hearer and therefore cannot be marked as indefinite (SAG, vol. 4, s.72). In OS, sentence-initial EN-subjects introduce text-referents of long lifespan (measured in number of co-referring further mentions), while non-initial EN-subjects introduce text-reference whose lifespan is shorter (there are fewer co-referring further mentions). This is shown in Table 5 below.

| | Period I | Period II | Period III |
|---|----------|-----------|------------|
| Co-referring further mentions after initial EN-subjects | 6,7 | 6,2 | 2,3 |
| Co-referring further mentions after non-initial EN-subjects | 2,3 | 4,0 | 2,4 |

Table 5. *Lifespan of EN-subjects in OS (average number of co-referring further mentions in the discourse)*

Table 5 illustrates that the presentative function, the first stage of grammaticalization of EN into indefinite article, is in period I, at the inception of the grammaticalization, fulfilled by not just any EN-phrase but mainly by sentence-initial EN-subjects. It is cataphoric, with numerous co-referring expressions following. In period II, all EN-subjects have a strong cataphoric function, although the sentence-initial ones still introduce more salient text-referents. In period III the position of EN-subject no longer influences the salience of the text-referent it introduces.

An example of the difference between the sentence-initial and non-initial EN-subjects is given below. The example comes from Jart, period II religious narrative.

(5) **EN diäkne** war j enom stadh som heet montepessolanus hulkin som foor mz diäwlskap oc truldom. thän sami for annars sins kompan's bö'n. the wilde wita tidhinde aff sino lande kalladhe til sin diäfwlen. ok diäfwlen vppinbaradhis j enna quinno liknilse ok syntis hanum idhkelika j tholiko liknilse Diäknin iäfwadhe hwat thät war häldir diäfwl hällas ey. ok tho spordhe han at diäwlin sit ärande ok han swaradhe hanum. Nu j the stundinne for ther fram vm **en prästir** mz gud's likama til *en siukan man* ok klokkan ringde for gud's likama. Ok ginstan wiste diäfwlen at thät war sandir gudh som prästin fördhe Ok ginstan j samu stundinne fiol han nidhir a iordhina a knä ok negh gud's likama Diäknin sa thetta ok sagdhe Nu tro iak sannelika at thz är ey diäfwl som thu kalladhe, for thy at han hafdhe ey fallit a knä som han giordhe Tha swaradhe diäfwlen ok sagdhe hanum. West thu ey hwat scriwat star j the hælgho skript som sighir J ihesu christi hedhir ok namn skulu böghias al knä. badhe j himerike ok iordhrike ok hælwite ok ginstan war han borto fran them

(Jart 3)

bara efter någon dag. 'The foot was worse. A smell started to be felt just after one day'. (Larsson 2006:205). In the second sentence the indefinite NP *en lukt* 'a smell' is anchored by the previous description of an illness, resulting in gangrene. The OS examples discussed here lack such anchoring.

‘There was a deacon in a town called Montpellier, who practiced witchcraft, he called the devil asked by his friend. They (i.e. deacon and his friend) wanted to know news from their country. And the devil revealed himself in a woman’s image and always appeared in such image. The deacon hesitated whether it was the devil or not and asked what he (the deacon) wanted of him (the devil) and he answered him. In this moment a priest went by with God’s body (travelling) to a sick man and the bell rang for the God’s body. And immediately the devil knew it was the true God that the priest carried. And immediately in the same moment he fell down on earth on his knees and bend for the body of God. The deacon saw this and said: Now I truly believe you are not the devil as you claimed, because he would not have fallen on his knees. Then the devil answered and told him: Do you not know what is written in the holy script which says that in the name of Jesus shall each knee bend, both in heaven and on earth and in hell. And immediately he was gone from them.’

There are two characters in the story: the deacon (*diäkne*), who is the main hero of the story, and the priest (*prästir*), who is a secondary, background character. While the first is introduced by means of a sentence-initial (and actually also text-initial) EN-subject, the background character is presented by means of a non-initial one. We may note further that there is one more character presented by means of an EN-phrase, *en siukan man* ‘a sick man’, introduced by means of a prepositional phrase and not picked up at all. The function of the EN-phrase (subject, object, PP) as well as its position seem to play an important role in how text-referents of different salience are introduced.

7. WHY THE PRESENTATIVE FUNCTION?

Why is the occurrence in the presentative function the first stage of grammaticalization of the indefinite article? We have noted that the presentative function is in fact cataphoric, since the EN-phrase is co-referring with expressions in later discourse (this is not necessarily true of the specific EN, since the new text-referent need not be picked up in later discourse). Interestingly enough, EN starts to be used in this function first when the anaphors are obligatorily marked.

In the oldest OS texts in period I (mainly AVL, but even younger legal texts) anaphors are marked only occasionally. The marking has a twofold character: accessible referents (accessibility is here understood as the distance between the antecedent and the anaphor measured in number of intervening words and the number of intervening syntagms) may be marked by preposed demonstrative *þæn* ‘this’ while the less accessible referents may be marked by means of the suffix -IN, originally a demonstrative *hinn* ‘yon’. Neither of these is obligatory in period I. However, in period II, the suffix has gained ground in the anaphoric contexts and the anaphors are obligatorily marked (Skrzypek 2012:102-107). In the example above the co-referring NPs are definite (*diäknin* ‘the deacon’ and *prästin* ‘the priest’). This is also the time when EN

starts occurring in the presentative function. This suggests an interdependence between the two developments. The presentative marking seems to develop in response to the already developed anaphoric marking, creating a symmetry – salient, persistent text-referents are marked as such already in the first mention (by EN) and in all further mentions, if they take form of an NP (by -IN). The development of regular, obligatory anaphoric marking is conducive to the development of the presentative marker. However, such a development does not take place in closely related languages, e.g. Icelandic (which to this day lacks an indefinite article).

8. CONCLUSIONS. A TYPOLOGICAL PERSPECTIVE

Grammaticalization of the indefinite article seems to always follow that of the definite article. Such is the case in other Indo-European languages at least, e.g. English (Hopper and Martin 1987), Spanish (Pozas Loyo 2010) or Italian (Stark 2002). The inception of the grammaticalization of the indefinite is always located in the purely textual domain: as a presentative device, particularly used with text-referents of some prominence. These are the text-referents which once introduced would in the wider discourse be marked as definite (anaphora). Since the presentative marker only marks salient (prominent) text-referents, its presence is cataphoric, obligatorily co-referring with an expression in later text. The embryonic articles are thus used mainly to track text-referents in the discourse and their original functions are textual.

We noted earlier that the development in Old English involved two elements, *sum* and *an*, which were used complementarily: *sum* to introduce new and salient text-referents, while *an* to introduce new but less salient text-referents. The former was moreover located at the beginning of an utterance, while the latter occurred later in the text. We do not see a similar pattern in Swedish, since there is only one form (a competing item is *någon* ‘some’, but it competes against EN in non-specific contexts and mainly in period 3, see Skrzypek 2012:184-190). We do see however that the position of EN plays a great role in its interpretation as presentative or specific marker.

The situation reported for Old Tuscan (Stark 2002) shows a number of forms competing in the presentative function (similarly to English, differently from Swedish), *uno* ‘one/a’ and *certo* ‘certain’ with specific reference and *alcuno* ‘nobody/none’, probably with non-specific reference (Stark 2002:325). Both *uno* and *certo* occur with important text-referents, while *alcuno* is used with non-specific interpretation (Stark 2002:330). The difference between the development in Old Tuscan and this in Old Swedish is such that in OS there is no form competing with EN in the presentative/specific contexts. A possible candidate is *viss* ‘certain’, but it does not seem to function in this way. More

interestingly, the non-specific *någon* can occasionally be used in specific contexts signalling that the new text-referent is unimportant, backgrounded and will not be picked up in later discourse (a function fulfilled by *någon* in Modern Swedish as well, see Nivre 2002). The position of the NPs introducing new text-referents in Old Tuscan has not been taken into account in the study, but none of the examples given shows text-initial ones in a pattern similar to the OS one.

In her study of the grammaticalization of the indefinite article in Spanish, Pozas Loyo states that the first non-numerical uses of *uno* were to introduce new and salient text-referents, an introduction which triggered a chain of anaphoric mentions (Pozas Loyo 2010:282). There is no direct mention of a regular and obligatory anaphoric marking in the book, however, the definite article in Spanish is of an earlier date than the indefinite, thus Spanish seems to follow the same pattern as Swedish, English and Tuscan. And again, there is a competing item, *algún*, which is used with non-specific referents and thus displays a complementary distribution with *uno* at the inception of grammaticalization of the latter.

It would be of great interest to include other, also not Indo-European, languages, in a study of the relative chronology of development of the definite and indefinite articles. The language sample given in WALS shows 194 possible candidates of languages with both definite and indefinite articles. Of equal interest is the study of languages with indefinite articles only (45 according to WALS) to see under what circumstances they arose, if their inception was also the presentative function and what referent tracking devices were used in the language at that time.

SOURCES

- AVL = Äldre Västgötalagen. In: *Samling af Sweriges gamla lagar*, ed. Hans Samuel Collin and Carl Johan Schlyter, 1827, vol. 1. Stockholm: Haeggström.
- Bo = *Bonaventuras betraktelser över Kristi liv*, ed. Gustaf Edvard Klemming, 1860. Samlingar utgivna av Svenska fornscriftsällskapet 15. Stockholm: Norstedt.
- Bur = Codex Bureanus. In: *Ett fornscrvenskt legendarium*, ed. Georg Stephens, 1847. Samlingar utgivna av Svenska fornscriftsällskapet 7:1. Stockholm: Norstedt.
- Did = *Sagan om Didrik af Bern*, ed. Gunnar Olof Hyllén-Cavallius, 1850-1854. Samlingar utgivna av Svenska fornscriftsällskapet 10. Stockholm: Norstedt.
- DL = Dalalagen. In: *Samling af Sweriges gamla lagar*, ed. Hans Samuel Collin and Carl Johan Schlyter, 1827, vol. 5. Stockholm: Haeggström.
- HML = Helga manna lefverne. In: *Klosterläsning*, ed. Gustaf Edvard Klemming, 1877-1878. Samlingar utgivna av Svenska fornscriftsällskapet 22. Stockholm: Norstedt.
- Jart = Järteckensbok. In: *Klosterläsning*, ed. Gustaf Edvard Klemming, 1877-1878. Samlingar utgivna av Svenska fornscriftsällskapet 22. Stockholm: Norstedt.
- JB = *Stockholms stads Jordebok 1420-1474*, ed. Hans Hildebrand, 1876. Kongl. Samfundet för utgifvande af handskrifter rörande Skandinaviens historia med understöd af Stockholms stad. Stockholm: Ivar Häggströms boktryckeri.

- KM = Karl Magnus. In: *Prosadikter från Sveriges medeltid*, ed. Gustaf Edvard Klemming, 1887-1889. Samlingar utgivna av Svenska fornscriftsällskapet 28. Stockholm: Norstedt.
- Linc JB = Själens kloster. In: *Skrifter till läsning för klosterfolk*, ed. F.A. Dahlgren, 1875. Samlingar utgivna av Svenska fornscriftsällskapet 20. Stockholm: Norstedt.
- Linc NR = Legendan om Sankt Amalberga. In: *Ett fornscrivenskt legendarium*, ed. Georg Stephens, 1847. Samlingar utgivna av Svenska fornscriftsällskapet 7:3. Stockholm: Norstedt.
- OgL = Östgöotalagen. In: *Samling af Sveriges gamla lagar*, ed. Hans Samuel Collin and Carl Johan Schlyter, 1827, vol. 2. Stockholm: Haeggström.
- Pent = *Fem Moseböcker på fornscrivenska* enligt Cod. Holm. A 1, ed. Thorell, Olof. Samlingar utgivna av Svenska fornscriftsällskapet 60. Uppsala: Almqvist & Wiksell.
- PK = Sveriges krönika (Prosaiska krönikan). In: *Småstycken på fornscrivenska*, ed. Gustaf Edvard Klemming, 1868-1881. Stockholm: Norstedt.
- SVM = Sju vise mästare. In: *Prosa dikter från Sveriges medeltid*, ed. Gustaf Edvard Klemming, 1887-1889. Samlingar utgivna av Svenska fornscriftsällskapet 28. Stockholm: Norstedt.
- Troj = *Historia Trojana*, ed. Robert Geete, 1892. Samlingar utgivna av Svenska fornscriftsällskapet 29. Stockholm: Norstedt.
- YVL = Yngre Västgöotalagen. In: *Samling af Sveriges gamla lagar*, ed. Hans Samuel Collin and Carl Johan Schlyter, 1827, vol. 1. Stockholm: Haeggström.
- All source texts apart from JB available also from *Fornscrivenska textbanken*, www.lu.se/nordlund.

Modern Swedish short stories

Ekelund, Fredrick. 2006. *Över gränsen*. Bonniers.

Jungstedt, Mari. 2006. En helg på landet. *Över gränsen*. Bonniers.

Larsson, Åsa. 2006. Systrarna Hietala. *Över gränsen*. Bonniers.

REFERENCES

- Barwise Jon & Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4, pp. 159-219.
- Brandtler, Johan and Lars-Olof Delsing. 2010. Framväxten av obestämd artikel i svenska. Definitheit, specificitet och polaritet. In: Byrman, Gunilla and Anna Gustaffson and Henrik Rahm. *Svensson och svenskan. Med sinnen känsliga för språk. Festschrift till Jan Svensson den 24 januari 2010*. Lund.
- Givón, Talmy. 1981. On the development of the numeral 'one' as an indefinite marker. *Folia Linguistica Historica* II, pp. 35-53.
- Gunnarsson, Britt-Louise. 1982. *Lagtexters begriplighet: en språkfunktionell studie av medbestämmandelagen*. Stockholm: LiberFörlag.
- Haskå, Inger. 1972. *Studier över bestämdhet i attributförsedda nominalfraser: uttrycksptyperna det vida hav, vida havet osv. i svensk poesi 1500-1940*. Lund: Studentlitteratur.
- Heim, Irene. 1988. *The Semantics of Definite and Indefinite Noun Phrases*. NY & London: Garland Publishing, Inc.
- Heine, Bernd. 1997. *Cognitive Foundations of Grammar*. Oxford: OUP.
- Heine, Bernd and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Hellquist, Elof. 1957. *Svensk etymologisk ordbok*. Lund: Gleerup.
- Hirvonen, Ilkka. 1997. *Konstruktionsstyperna gamle man, den gamle man och den gamle mannen. Studier i nordisk filologi*. Svenska litteratursällskapet i Finland: Helsinki.

- Hopper, Paul J. and Janice Martin. 1987. Structuralism and diachrony: The development of the indefinite article in English. In: Ramat, Anna Giacalone, Onofrio Carruba & Giuliano Bernini. *Papers from the Seventh International Conference on Historical Linguistics*, pp. 295-304.
- Håkansson, David. 2008. *Syntaktisk variation och förändring. En studie av subjektslösa satser i fornsvenska*. Lund: Lunds universitet (Lundastudier i nordisk språkvetenskap A:64).
- Nivre, Joakim. 2002. Three Perspectives on Swedish Indefinite Determiners. *Nordic Journal of Linguistics* 25, pp. 3-47.
- Pozas Loyo, Julia. 2010. *The development of the indefinite article in Medieval and Golden-Age Spanish*. University of London.
- SAG = Teleman, Ulf, Staffan Hellberg and Erik Andersson (eds.) 1999. *Svenska Akademiens grammatik* vol. 1-4. Stockholm: Norstedts ordbok.
- Skrzypek, Dominika. 2012. Grammaticalization of (in)definiteness in Swedish. Poznań: Wydawnictwo Naukowe UAM.
- Stark, Elisabeth. 2002. Indefiniteness and specificity in Old Italian texts. *Journal of Semantics* 19, pp. 315-332.
- Syrett, Martin. 2002. Morphological developments from Ancient Nordic to Old Nordic. In: Bandle, O. et. al (eds.), *The Nordic Languages*. Berlin/New York: Walter de Gruyter, pp. 719-727.
- WALS = The World Atlas of Language Structures, wals.info

Dominika Skrzypek

Uniwersytet im. Adama Mickiewicza w Poznaniu
Katedra Skandynawistyki
Collegium Novum
al. Niepodległości 4
61-874 Poznań
Poland
dosk@amu.edu.pl